HIGHLAND STATISTICS COURSES
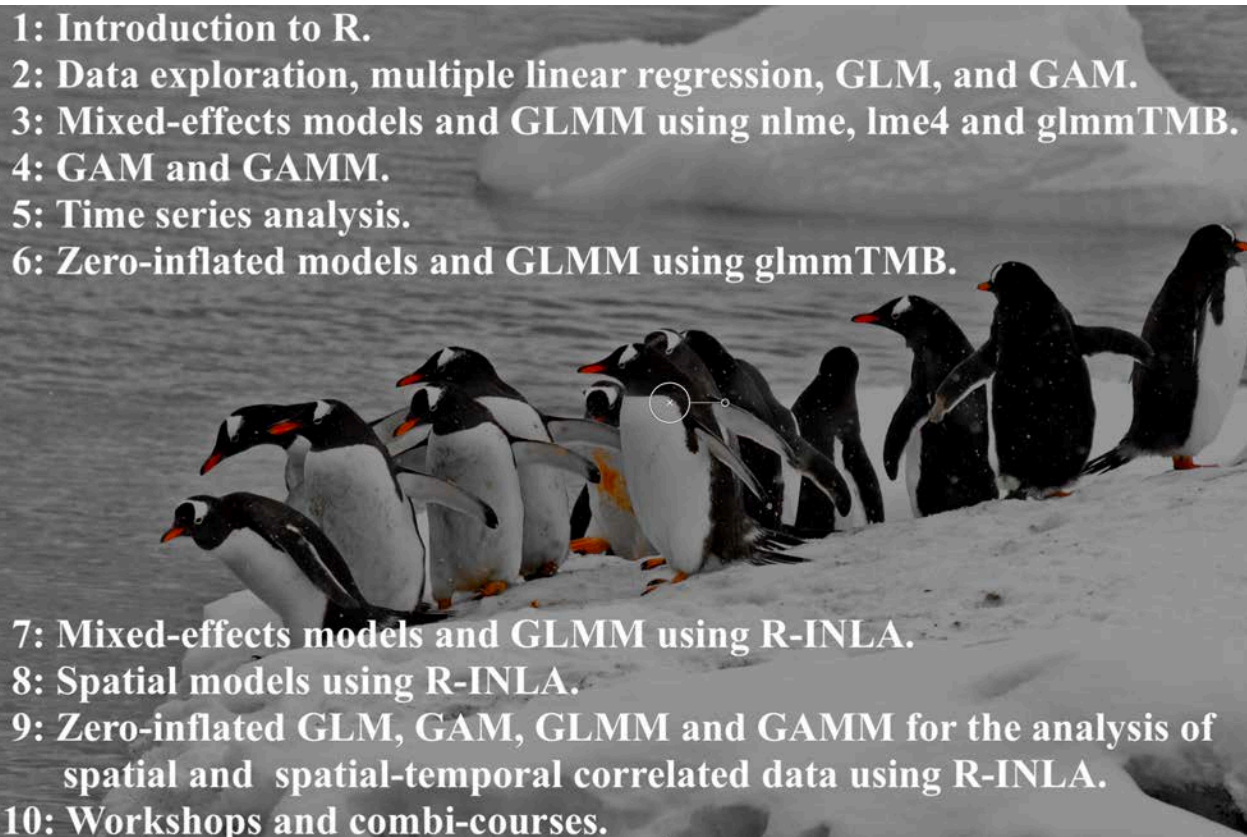
1: Introduction to R.
2: Data exploration, multiple linear regression, GLM, and GAM.
3: Mixed-effects models and GLMM using nlme, lme4 and glmmTMB.
4: GAM and GAMM.
5: Time series analysis.
6: Zero-inflated models and GLMM using glmmTMB.
7: Mixed-effects models and GLMM using R-INLA.
8: Spatial models using R-INLA.
9: Zero-inflated GLM, GAM, GLMM and GAMM for the analysis of spatial and spatial-temporal correlated data using R-INLA.
10: Workshops and combi-courses.

Highland Statistics Ltd. provides 10 different statistics courses. The course instructors are Dr. Alain Zuur (statistician) and Dr. Elena Ieno (biologist). Our statistical and biological backgrounds ensure a lively and enjoyable interaction with our participants. One of our strong points is explaining statistics in a non-technical and understandable language.

Some of our (online or onsite) courses are run as in-house courses whereas other courses are open. For a list of upcoming courses, see: http://highstat.com.

If you have various colleagues who are interested in one of our courses it may be more cost-effective to organise a course at your institute. This can be done as an in-house course or as an open course. With an in-house course, you decide who participates and we charge a fixed fee. Our fee will depend on the country. For an open course, we will require a conference room (plus projector) for about 30 people, and an additional 10 local participants.

SHORT COURSE DESCRIPTION

Course 1: Introduction to R using a protocol for conducting and presenting results of regression-type analyses

- In this course, we provide an introduction to R and at the same time explain how to conduct data exploration, apply (simple) linear regression models, communicate results, and also determine optimal sample size (using power analysis) in case you want to set up a new field study or experiment.
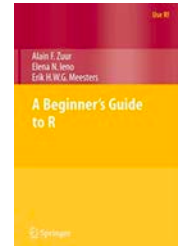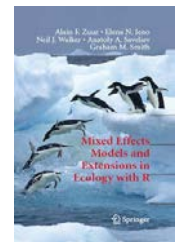
Course 2: Data exploration, regression, GLM and GAM: with an intro to R.

- We begin with an introduction to R and provide a protocol for data exploration to avoid common statistical problems. We will discuss how to detect outliers, deal with collinearity and transformations. An important statistical tool is multiple linear regression. Various basic linear regression topics will be explained from a biological point of view. We will discuss potential problems and show how generalised linear models (GLM) can be used to analyse count data, presence-absence data and proportional data. Sometimes, parametric models (linear regression, GLM) do not quite fit the data and in such cases generalised additive models (GAM; a smoothing technique) can be used.
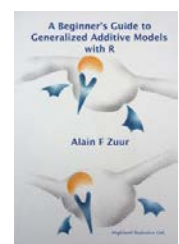
Course 3: Introduction to mixed effects models and GLMM

- The course starts with a short revision of multiple linear regression and generalised linear models, followed by an introduction to linear mixed-effects models and generalised linear mixed-effects models (GLMM) to analyse hierarchical or clustered data, e.g. multiple observations from the same animal, site, area, nest, patient, hospital, vessel, lake, hive, transect, etc. In the second part of the course GLMMs are applied on continuous (e.g. biomass), binary (e.g. absence/presence of a disease), proportional (e.g. % coverage) and count data using the Gaussian, Poisson, negative binomial, Bernoulli, binomial, beta, and gamma distributions.

Course 4: Introduction to GAM and GAMM

- We start with a short revision of data exploration and linear regression. We then introduce generalised additive models (GAM) to model non-linear relationships. We will execute these models in mgcv. In Module 2, we will revise linear mixed-effects models and show how to implement a generalised additive mixed-effects model (GAMM). We also show how to include an interaction between a smoother and a categorical covariate. In Module 3, we will revise basic GLMs and extend these towards GAMs. In Modules 4 and 5 we will discuss GAMMs for the analysis of count data,

absence-presence data, proportional data, and continuous data. We also discuss 2-dimensional smoothers (including the soap-film smoother for study areas with barriers; e.g. an island in the sea).

Course 5: Time series analysis using regression techniques

- The course starts with a short revision of data exploration and multiple linear regression models. A non-technical introduction of generalised additive models (GAM) is provided. GAMs will be used to estimate long-term trends, seasonal patterns, covariate effects and auto-regressive correlation. We also provide a short introduction to linear mixed-effects models and generalised linear mixed-effects models (GLMM) to analyse hierarchical data (e.g. short time series from the same core or site). GLMMs and GAMMs are used to estimated trends, seasonality and covariate effects in multivariate time series.

Course 6: Introduction to zero-inflated models

- The course starts with a short revision of data exploration, multiple linear regression and Poisson GLM. We then discuss 3 more models for the analysis of count data, namely the negative binomial, generalised Poisson and Conway-Maxwell-Poisson GLMs. After a short theory presentation in which we explain how to extend these models towards zero-inflated models, we apply them to various data sets. We also use the Tweedie GLM and the zero-altered Gamma GLM for the analysis of zero-inflated continuous data. In the second part of the course, we start with a short revision of linear mixed-effects models. This is followed by a series of exercises in which we analyse zero-inflated count data, continuous data, and proportional data using zero-inflated GLMMs. Throughout the course we will use the glmmTMB package in R.

Course 7: Mixed-effects models and GLMM using R-INLA

- The course begins with a brief revision of multiple linear regression, followed by an introduction to Bayesian analysis and how to execute regression models in R-INLA. We then explain linear mixed-effects models to analyse nested data, followed by a series of mixed modelling exercises in R-INLA. Nested data means multiple observations from the same animal, site, area, nest, patient, hospital, vessel, lake, hive, transect, etc. In the second part of the course GLMMs are applied on count data, binary data (e.g. absence/presence of a disease), proportional data (e.g. % coverage) and continuous data (e.g. biomass or distance) using the Poisson, negative binomial, Bernoulli, binomial, beta and gamma distributions.

Course 8: Intro to Regression Models with Spatial Correlation using R-INLA

- We begin with an introduction how to add dependency to regression models using frequentist tools. After discussing the limitations of this approach we switch to Bayesian

techniques. R-INLA is used to implement regression models, generalised linear models (GLM) and generalised linear mixed-effects models (GLMM) with spatial dependency

Course 9: Zero-inflated GLM, GAM, GLMM and GAMM for the analysis of spatial and spatial-temporal correlated data using R-INLA

- We will start with a short revision of multiple linear regression, followed by a basic introduction to Bayesian analysis, and we show how to execute a linear regression model in R-INLA. In the second module, we will explain how to deal with zero-inflated count data and zero-inflated continuous data using zero-inflated Poisson, zero-inflated negative binomial, and zero-inflated Gamma GLMs. We also explain hurdle models. In the third module, we will introduce generalised additive models (GAM) to model non-linear relationships. We show how to execute these in mgcv and also in R-INLA. In the fourth part of the course, we will revise linear mixed-effects models and implement these in R-INLA. We also apply generalised linear mixed-effects models (GLMM) and generalised additive mixed-effects model (GAMM) in R-INLA. In the fifth part of the course, we will apply zero-inflated GAMs and GAMMs (and GLMMs) on various spatial correlated data sets. In module 6, we apply GAM, GAMM, and GLMM on spatial-temporal correlated data. We also deal with natural barriers for the spatial correlation (e.g. benthic species that live on a coral reef around an island). We will use barrier models; these ensure that spatial correlation seeps around a barrier (in this case an island). All exercises are executed in R-INLA.
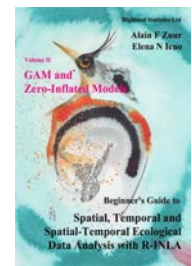
Course 7: Workshop and combi-course

- Combine the appropriate modules and use your own data sets during the course.

RECOMMENDED ORDER OF COURSES

If you do not have spatial or temporal data, then we recommend to attend the following courses within a time span of 3 years.

1. Introduction to data exploration, regression, GLM and GAM. With introduction to R (course 2).
2. Introduction to mixed effects models and GLMM (course 3).
3. Depending on whether you have zero-inflation and/or non-linear relationships and/or time series, you can then attend the 'Introduction to GAM and GAMM', 'Introduction to zero-inflated models', or the 'Time series' course (courses 4, 5 or 6).

If you have spatial (or spatial-temporal) data, then we recommend the following courses.

1. Intro to data exploration, regression, GLM and GAM. With introduction to R (course 2).
2. Introduction to mixed effects models and GLMM using R-INLA (course 7).
3. Introduction to Regression Models with Spatial Correlation using R-INLA (course 8).
4. If you have zero-inflation and/or non-linear relationships and spatial (-temporal) data, you can then attend the 'Zero-inflated GLM, GAM, GLMM and GAMM for the analysis of spatial and spatial-temporal correlated data using R-INLA' course.

It is possible to skip the 'Introduction to mixed effects models and GLMM using R-INLA' course before taking the spatial INLA courses, but this means a steeper learning curve.

GENERAL INFORMATION

All courses are non-technical, are taught in R, and use a course website that contains:
*   5 - 10 theory presentations (on-demand video and also downloadable pdf files).
*   15 - 25 exercises (on-demand video, with downloadable data sets and documented R code).
*   A Discussion Board where you can ask course-related questions.
*   Live chat facilities for short questions.
*   Recorded Zoom sessions from previous 'Live' online courses are available to watch.

Access to the course website is for 12 months.

Live interaction.
*   All courses include a 1-hour face-to-face video chat with the instructors. You can ask questions about your own data.
*   For the self-study courses, you can request 4 times a 30 minutes video chat with the instructors and ask questions related to the course material.
*   Live interaction needs to take place within 12 months after being given access to the course website.

*   You can start an online course at any time.
*   The workload of each course is about 40 hours.
*   A certificate (pdf file) will be provided upon completing the course or after completing a short online assessment.

For further information go to: http://highstat.com

Email: Dr. Alain F. Zuur at highstat@highstat.com